5-9-2018

Lessons from my Undergraduate Research Students

Paul A. Craig
Rochester Institute of Technology

Recommended Citation

Craig, P. A. (2018) Lessons from my undergraduate research students. J Biol Chem. 293(27) 10447–10452 DOI: 10.1074/jbc.RA118.003722

JBC ASBMB AWARD ARTICLE

Lessons from my undergraduate research students

Published, Papers in Press, May 9, 2018, DOI 10.1074/jbc.RA118.003722

Paul A. Craig[1]

From the School of Chemistry and Material Science, Rochester Institute of Technology, Rochester, New York 14623

Edited by Norma M. Allewell

From very early on, my personal/professional life has been shaped by teachers in many different settings. Teaching and learning form a two-way street. In the process of teaching undergraduate students, particularly in the research lab, I have learned some profound lessons about the importance of listening to them, challenging them, giving them autonomy, and allowing them to enjoy success and to risk failure. I am now working with a team of faculty members to implement these lessons in a course-based undergraduate research experience in the biochemistry teaching laboratory. Our goal is to seek answers to the question "How do students become scientists?" and to implement those answers with our future students.

I am going to begin by talking a little bit about my journey. Then, I will focus on the joys of working with students, particularly undergraduate students, their scientific achievements, and the lessons they have taught me.

I have been blessed throughout my life to work with people who challenged me and loved me. In eighth grade, Mr. Mayo told me that I would make a good scientist someday if I would ever learn to follow directions. In high school, Mr. Johnstone, Mrs. Radke, and Mr. Joyal taught me the beauty of chemistry–where numbers and matter merge to provide remarkable explanation for things like stoichiometry and redox reactions, and eventually introducing me to the beauty of organic chemistry and biochemistry. In college, Dr. Dale Williams challenged me to look deeper into the numbers and substance of chemistry. In graduate school, Dr. Dekker taught me to methodically explore natural processes, documenting and disseminating my findings to my community. When I entered the faculty ranks, my department head at Rochester Institute of Technology (RIT), Jerry Takacs, found opportunities for me to combine my love for biochemistry with my interest in computers. Chris Rohlman shared a copy of his successful National Science Foundation (NSF) DUE proposal, which helped me to win my first external funding. The first time I presented an educational effort that combined computers and biochemistry, Judy Voet stopped by my poster and offered me great encouragement: "We need to see more of this." I shall never forget the first Biochemistry Education Symposium in San Francisco a few years later. We had so much we wanted to share with each other that there wasn't time for lunch one day. So, during a noon poster session, I saw Judy Voet walking around the poster session with apples and granola bars for everyone who wanted them! In 1999, shortly after I received tenure, Dick Doolittle, my colleague at RIT, saw that I was on the verge of burnout and told me, "You've got to get out of here," so I started looking for a sabbatical. In 2001, Phil Bourne, who was then directing the computational efforts of the RCSB Protein Data Bank, invited me to spend a sabbatical year with him, where I had the chance to see how the power of computers could be harnessed to serve the growing data needs of our community. In 2005, with support from the NSF, my students and I spent the summer at Brookhaven National Laboratory, where I witnessed a productive scientific community and where I also met my extraordinary collaborator, Herbert Bernstein, who taught me how to become funded and remain funded. I could mention many more scientists who have been giants in my life.

In addition to talking about those who have taught me science, I must also talk about those who taught me faith, for I am both a scientist and a Christian. I am and have been surrounded by a great cloud of witnesses who have challenged and encouraged me to engage in the lifelong pursuit of faith. Billy Washington taught me to ask any question that I have. Helen Crabtree taught me to dig deeper into the Bible and look for connections. Louise Church spoke to me about integrity. She also taught me that public speaking is not the same thing as reading. Dr. Dale Williams at Oral Roberts University encouraged me to pursue the path that lay before me at a time when I was thinking of pursuing a career outside of science. Others have spoken words of encouragement during dark days, particularly my mentor in graduate school, Dr. Eugene Dekker. Bob Bateman has been my friend and confidant for many years, starting with a conversation at Ann Arbor in 1984, as we have walked the journey of science and faith together. These people have taught me the importance of respecting others and building personal integrity in all of life, which is best encapsulated in a saying that I learned from Steve Murray when he was pastor at La Jolla Presbyterian Church during my sabbatical with Phil Bourne:

- Sew a thought, reap an action
- Sew an action, reap a habit
- Sew a habit, reap a character
- Sew a character, reap a destiny

I have had the privilege of working with many wonderful people throughout my career. One of the first opportunities

This work was supported by Rochester Institute of Technology, NIGMS, National Institutes of Health Grant R15GM078077, and by National Science Foundation Grants 0402408, 1503811, and 1709170. The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation.

[1] To whom correspondence should be addressed: Rochester Institute of Technology, 1 Lomb Memorial Dr., Gosnell Bldg., Rm. 2116, Rochester, NY 14623-5603. Tel.: 585-475-6145; E-mail: paul.craig@rit.edu.

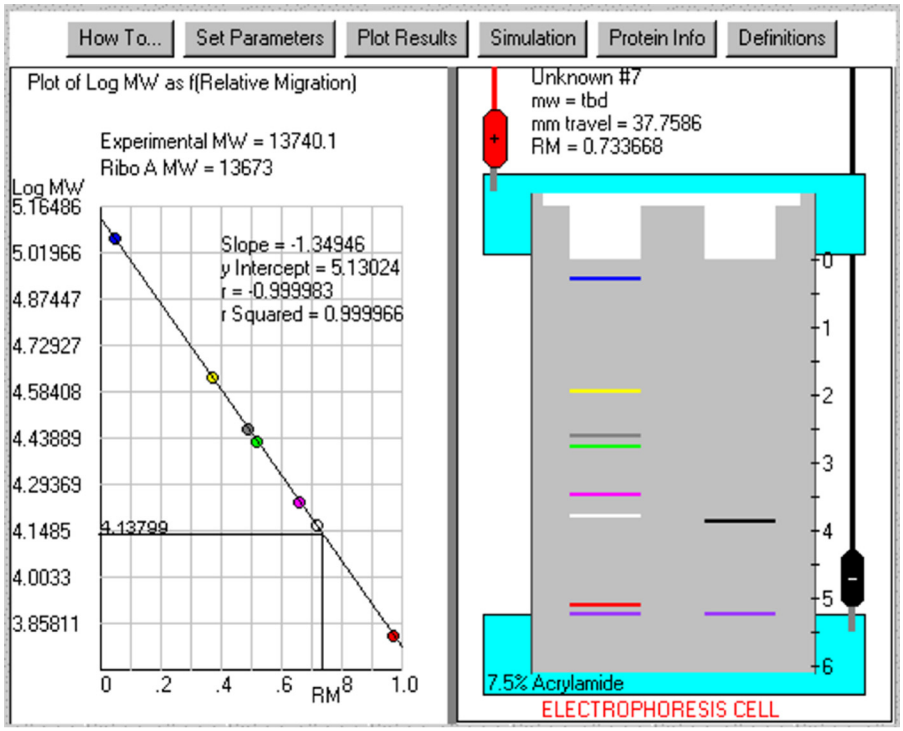

Figure 1. Simulation of 1D electrophoresis of proteins.

was a group effort on a digital library for biochemistry called Biomolecules-Alive (1). It was not a success, but it was a great learning experience with terrific people. I mention this project to point out that failure is part of the academic life, and it can occur during projects, and it frequently does occur with grant proposals and manuscript submissions. I encourage students when they encounter failure to consider the quote from Teddy Roosevelt, who spoke of the person who "fails while daring greatly" (2). I also encourage faculty to be honest with their students about their own failures and how they responded at those times. In 2010, Bob Bateman and I published a rubric about assessing student molecular literacy for the spring 2010 PDB Newsletter (3). This led to a talk about assessing student learning with molecular visualization at the 2013 ASBMB-sponsored symposium on Student-Centered Education in the Molecular Life Sciences. The BioMolViz group formed at that time and is expanding this rubric to include systematic learning goals and objectives (4). If you use molecular visualization in your teaching, I strongly encourage you to participate in one of the BioMolViz workshops.

Perhaps the best part of this journey has been the honor of working with gifted young people. From this point forward, I will focus on what they have accomplished, and the lessons they have taught me. The following stories mark highlights from the past 25 years, in which I hope to communicate my admiration and gratitude to a number of the young scientists who have been part of my life.

Research projects and lessons

Lesson #1: Students will make remarkable progress if you give them some freedom

David Mix is a B.S. chemistry graduate of RIT and was working on his M.S. in Computer Science. We did a project together in a graphics computer course, and he approached me to do a project for credit and I responded: "What would you like to do?" The Internet was quite new at the time, and I had heard about computational animations in web pages. I told him I wanted a killer web page that simulated electrophoresis of proteins. We started from an image in the Bio-Rad catalogue, and he created an application that many of you may have tried. It formed the basis for a whole string of projects that have involved more than 25 students since 1998. Fig. 1 is a screen capture image of the simulated protein separation in 1DE.

Lesson #2: Students will surprise you if you give them a challenge

In 1999, Janine Garnham was an undergraduate biochemistry major with an interest in computer science, who asked me if I had a biochemistry problem she could solve as a project in her computer science course. I was considering expanding the 1DE simulation to 2DE, so I asked her to create code that would predict the pI and molecular weight for a protein; so she did. Next, I met Jill Zapoticznyj, an undergraduate computer science major, who wanted to program at the interface of computer science and biology, so I challenged her to take Janine's code and build an application with a graphical user interface. Over the course of 3 years and multiple iterations, she completed the project, which I first presented at the ASBMB conference in 2004. Fast-forwarding to the fall of 2008, a first year student named Amanda Fisher was one of my advisees. We enjoyed talking together, and she wanted to join my research group, but at that time I had no projects that seemed to fit her interests. She visited me one last time before going home for the summer, and I told her that I had a great project that she could join but she would need to learn Java. So she did. Over the next 3

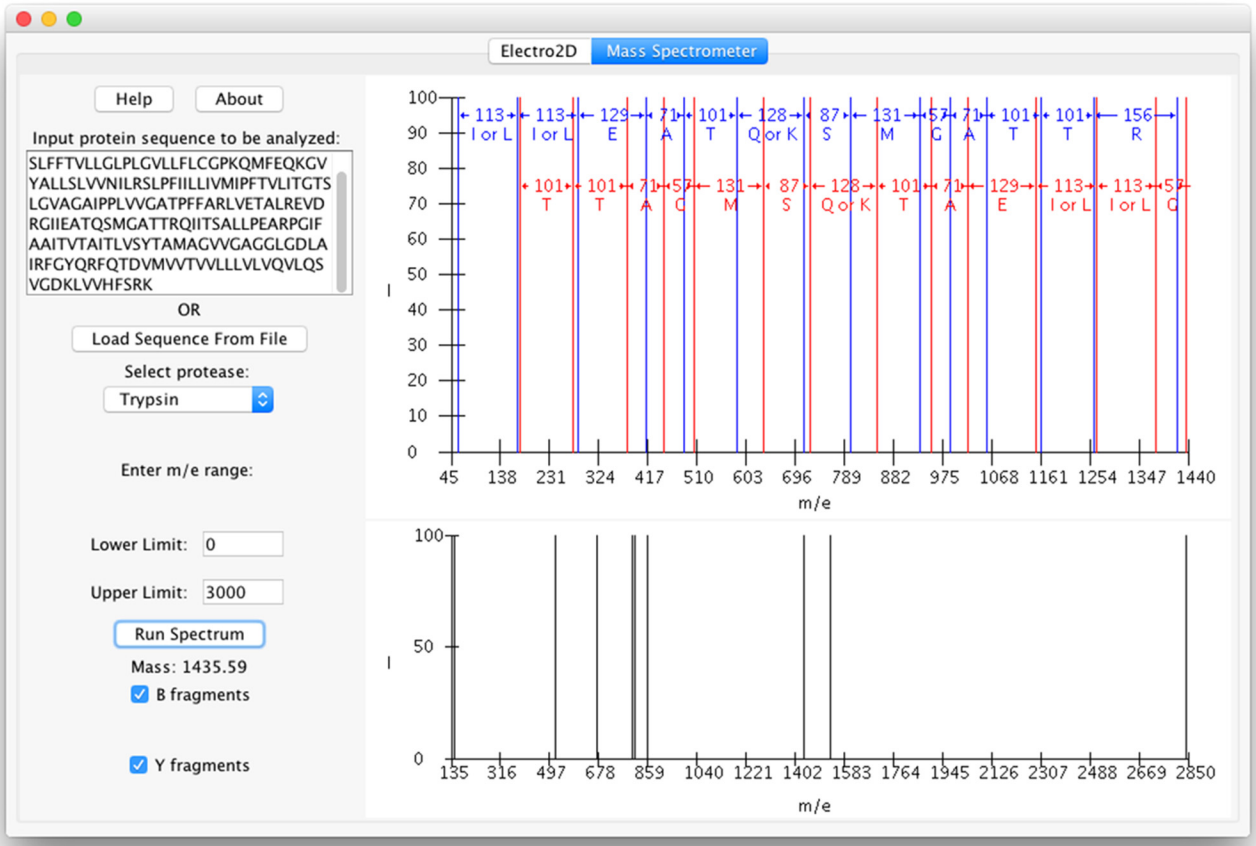

Figure 2. 2DE-tandem MS simulation with bacterial proteins.

years, we had a wonderful adventure together as she refined our existing code and interface and then elected to add a simulated tandem mass spectrometer, which I have used in biochemistry and analytical chemistry courses since then (Fig. 2) (5).

Lesson #3: Students sometimes know best

From this point on, I will focus on a project that began in our research group and eventually formed the basis for a very exciting Course-based Undergraduate Research Experience. Several students in my biochemistry class in 2004 really enjoyed a molecular visualization project, so I invited them to join an independent study on 3D molecular visualization, using a stereo projection system we had just obtained with support from the NSF ATE program. We were looking for a molecular visualization program that supported stereo graphics. One student did not stand out in the classes he took with me, but he was into video games and graphics. He found PyMOL, which had built-in support for stereo projection, and really kick-started the project. Two other students, Chris Parkin and Laura Grell, were bioinformatics majors who wanted to continue with the project after the semester, so the same NSF ATE grant supported them for a full summer at Brookhaven National Laboratory. I suggested that they write some code for PyMOL, to help users overcome the challenging command line interface, and I showed them where the tools they developed would fit on the PyMOL interface. I then had to leave for a family obligation and told them to "do what they thought best." Well, they thought that my suggestion to work within the PyMOL graphical user interface was not the best idea, and they decided to build a plugin for PyMOL called EZ-Viz, which provided users with dropdown menus for selection and display options and even a few simple movies (6).

Lesson #4: Students moved us from teaching to research

The following summer, I hired two biotechnology majors, Brett Hanson and Charlie Westin, to work at Brookhaven National Laboratory on the project. I gave them a list of 10 projects to implement with EZ-Viz over the course of the summer. They called me at the end of week 1 and said: "We've completed what you asked us to do. What's next?" I was at a loss and asked them what they wanted to do. They told me that they thought they could use the tools in PyMOL to do some protein alignments and function prediction. I warned them that they would need to learn to code in Python. Undaunted, they spent the next 9 weeks building ProMOL from EZ-Viz. Brett and Charlie would be the first to admit the code was buggy and hideous, but they were able to demonstrate active-site alignments between proteins. Subsequent students, most notably Mario Rosa, cleaned up the code to provide a more stable and useful interface that provides users with options to compare proteins to a library of active-site motifs, to create their own motifs, and to assess the quality of their results by Levenshtein

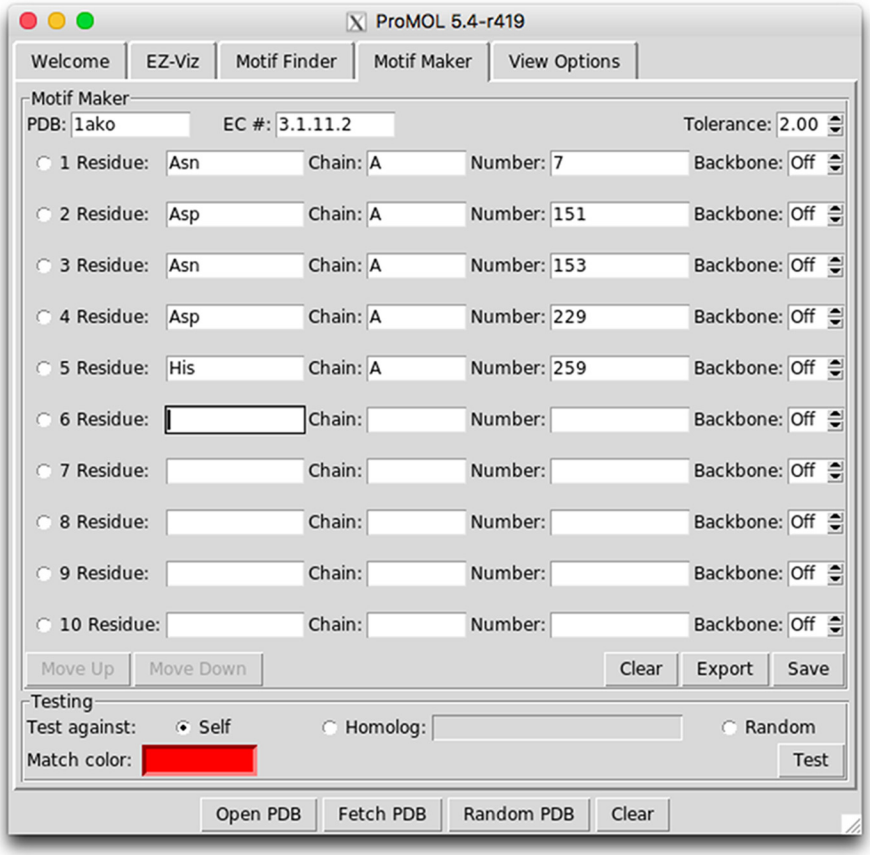

Figure 3. ProMOL interface for active-site motif creation. The five active-site residues from PDB entry 1AKO, an exonuclease from *E. coli*, were added to a motif template in the ProMOL library by residue name, chain, and residue number.

distance (7), root mean square deviation, and visual alignment in PyMOL (Fig. 3) (8).

Lesson #5: Given ownership of a challenging scientific problem, tools, and guidance, students start to think and act like scientists

To be honest, it took me several years to truly understand the power of the tool that these students had created and refined. Many students helped us to build the active-site motif library for this project (Fig. 4), drawn mainly from the Catalytic Site Atlas (9) provided by the European Bioinformatics Institute. At this point, we had a powerful tool looking for a problem to solve. The structural genomics initiative was in full swing, producing protein structures at an amazing rate, often with little or no annotation. My primary collaborator on this project, Herbert J. Bernstein, and I began to encourage our students to look into these proteins of unknown function. While I was fascinated by the beauty and predictive power of active-site alignments, our research students, notably Greg Dodge, started looking for other tools that predict protein function and developed a pipeline that integrated our results with ProMOL with results from BLAST (10), Pfam (11), and Dali (12). Using this approach, a small army of students queried every available "protein of unknown function" found in the Protein Data Bank with this array of tools, resulting in a publication of more than 50 high-quality alignments (13). I should also mention that Jeff Mills joined the faculty at RIT at this time and was key in helping us to understand and integrate Pfam with the project. While I was completely satisfied to focus only on the computational aspects of the project, these students, notably Greg Dodge, Enidza Nicole Arroyo, and Kaitlin Hart, insisted on going into the wet lab to test

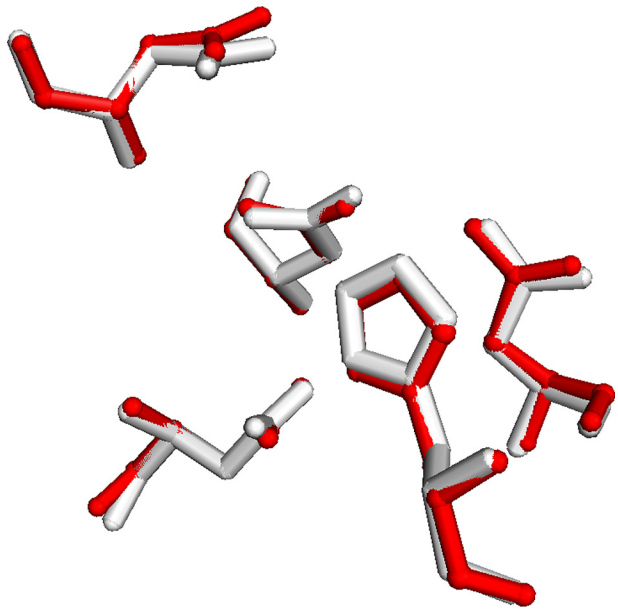

Figure 4. Alignment of PDB entry 3L1W (query) with 1AKO (template), an exonuclease from *E. coli* with ProMOL. The five residues in the alignment are Asn-7 (3L1W):7(1AKO), Asp-166:151, Asn-168:153, Asp-215:229, and His-242:259.

Table 1
The BASIL team

Institution	Name
California Polytechnic State University, San Luis Obispo, CA	Anya Goodman, Ashley Ringer-McDonald
Hope College	Mike Pikaart
Lagrange College	Arthur Sikora
Oral Roberts University	Bob Stewart
Purdue University	Trevor Anderson, Stefan Irby
RIT	Herbert Bernstein, Paul Craig, Jeff Mills, Suzanne O'Handley
St. Mary's University	Colette Daubner
SUNY Oswego	Webe Kadima, Julia Koeppe
Ursinus College	Rebecca Roberts

Table 2
The BASIL modules

In vitro modules	*In silico* modules
Protein expression	BLAST
Protein purification	Dali
Protein concentration	Pfam
SDS-PAGE	PyMOL and ProMOL
Enzyme activity	PyRx

their hypotheses. They expanded our pipeline to include *in vitro* characterization, where they obtained plasmids containing the genes for the proteins of interest, expressed and purified the proteins, identified and ordered substrates for the proteins, and then tested them for activity in the laboratory. These students hypothesized about the function of these proteins. In some cases they found activity; in others they did not. They experienced success and failure in the laboratory. They explored the literature. They contacted other scientists–in fact, one time Greg called the director of the Structural Genomics Initiative at the Argonne National Laboratory without even telling me. They had a great conversation. The students wrote about and presented their results.

The BASIL project

At this point, two things happened that led to what we now call the BASIL (Biochemistry Authentic Scientific Inquiry Lab) project. First, we began to notice a pattern–students on this project were making amazing strides as scientists. Second, and perhaps the more important catalytic event– our National Institutes of Health (NIH) funding dried up. Our initial support came from the NSF Advanced Technological Education program. We then had two rounds of funding through the NIH AREA (R15) mechanism. With the help of program officers from NIH and NSF, we started applying to the NSF IUSE Exploration and Design Tier. Our initial proposal included RIT and two other campuses and was not funded, but our program officer encouraged us strongly to recruit a stronger team and resubmit the proposal. So, Herbert Bernstein and I presented a poster at the 2014 ASBMB conference entitled, "Role of Undergraduate Biochemistry Education in Protein Function Assignment."

To summarize briefly, several people visited the poster and expressed interest in joining this effort to create a Course-based Undergraduate Research Experience (CURE) on eight campuses, with educational assessment from a ninth campus (Table 1). Since we began working together in July, 2015, we have created 10 lab modules (Table 2) with eight accompanying videos, taught the lab more than 20 times on our different campuses, presented our work at many conferences, worked out some of the rough edges on the project, created a blog with access to the modules and videos (14), and are in the process of assessing growth for the students in our courses, the teaching assistants in our labs, and the members of our faculty team.

As we implement the BASIL project, we have experienced the challenge of transitioning from traditional labs and inquiry-based labs to discovery-based labs, where neither the instructors nor the students know the answers to the question. There is a very distinct possibility that students may have success in some aspects of the course (*in silico* alignments, protein expression, and purification) but still not realize the ultimate goal of verifying the predicted protein function. Like all true research, the lab experience includes the very real possibility of success or failure.

Our overall goal in BASIL is to answer the question: "How do students become scientists?" We have made some progress. The instructors all consented to be interviewed about their experiences in the BASIL project. Trevor Anderson and Stefan Irby from Purdue University helped me prepare some questions that focused on the logistics of the project. The results are summarized here, and the full details have been published (15).

- There should be well-established protocols and uniform access to all of the software tools.
- It would be better to start with a known positive for the alignment studies and enzyme activity assays.
- Whereas the main material expense was for purchasing substrates, the real cost is the amount of time it takes for faculty members to implement the course: " . . . it cost me about 10 more hours per week," according to one of our team members.
- There has been a lot of camaraderie as we struggled to prepare the modules and to learn how to implement the computational tools on each campus.
- The BASIL team members have all been excited to present their results at meetings of the ASBMB, American Chemical Society, Biophysical Society, and Biennial Conference on Chemical Education, and had these things to say about it: "I hope this research approach will become the norm." "The issue is to highlight the idea that we are building scientists who are doing true science. There is a learning curve to implement this. We have to show the deep worth of it to newcomers."
- All team members expressed a strong desire to meet in a physical location. To date, we have held many video chat sessions, and different smaller groups of us have met at the conferences listed above.

We have had some interesting experiences with our students and have stories to tell about student growth, particularly during a joint poster session among students from two classes at Ursinus College, a traditional protein lab and a computational biochemistry course. The students appeared to learn a great deal from their interactions and expressed a desire to learn much more. However exciting this was, it was only anecdotal, and we are attempting to assess deeper questions as we grow.

1) Can we conduct research in a teaching lab?

2) How can we assess student progress as scientists?
3) What do we need to change about the way we teach to make this possible?

Future plans

- Create fully annotated faculty resources.
- Continue to explore our basic questions about student growth as scientists in the CURE setting.
- Recruit more campuses to participate.
- Create a model for student sharing of their results across campuses.
- Expand the project to new enzyme families.
- Build relationships with national databases for publication of student findings.

Conclusion

It has been my privilege to work with talented and generous colleagues throughout my career. My greatest joy as a professor has been to watch the personal and professional growth of my students. They have taught me a great deal, and I aspire to pass it on to future students, at least in part through the BASIL project.

Acknowledgments—I am deeply grateful to my former research students, the founding members of the BASIL collaboration (Herbert J. Bernstein, Susan Colette Daubner, Anya Goodman, Julia Koeppe, Suzanne F. O'Handley, Jeffrey L. Mills, Mike Pikaart, Ashley Ringer McDonald, and Rebecca Roberts), and the RCSB Protein Data Bank for providing structural data for all of the proteins that are studied in the BASIL project.

References

1. Craig, P. A. (2006) BioMoleculesAlive.org: The biochemistry and molecular biology digital library update. *Biochem. Mol. Biol. Educ.* **31,** 151–152 CrossRef
2. Roosevelt, T. (1910) The Man in the Arena, April 23, 1910, Theodore Roosevelt Speeches, Roosevelt Almanac
3. Bateman, R. C. and Craig, P. A. (2010) A proficiency rubric for biomacromolecular 3D literacy. *Protein Data Bank Newsl.* Number 45, 5–7
4. Dries, D. R., Dean, D. M., Listenberger, L. L., Novak, W. R., Franzen, M. A., and Craig, P. A. (2017) An expanded framework for biomolecular visualization in the classroom: learning goals and competencies. *Biochem. Mol. Biol. Educ.* **45,** 69–75 CrossRef Medline
5. Fisher, A., Sekera, E., Payne, J., and Craig, P. (2012) Simulation of two-dimensional electrophoresis and tandem mass spectrometry for teaching proteomics. *Biochem. Mol. Biol. Educ.* **40,** 393–399 CrossRef Medline
6. Grell, L., Parkin, C., Slatest, L., and Craig, P. A. (2006) Ez-Viz, a tool for simplifying molecular viewing in PyMOL. *Biochem. Mol. Biol. Educ.* **34,** 402–407 CrossRef Medline
7. Levenshtein, V. I. (1965) Binary codes capable of correcting deletions, insertions, and reversals. *Doklady Akademii Nauk SSSR* **163,** 845–848
8. Hanson, B., Westin, C., Rosa, M., Grier, A., Osipovitch, M., MacDonald, M. L., Dodge, G., Boli, P. M., Corwin, C. W., Kessler, H., McKay, T., Bernstein, H. J., and Craig, P. A. (2014) Estimation of protein function using template-based alignment of enzyme active sites. *BMC Bioinformatics* **15,** 87 CrossRef Medline
9. Furnham, N., Holliday, G. L., de Beer, T. A., Jacobsen, J. O., Pearson, W. R., and Thornton, J. M. (2014) The Catalytic Site Atlas 2.0: cataloging catalytic sites and residues identified in enzymes. *Nucleic Acids Res.* **42,** D485–D489 CrossRef Medline
10. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402 CrossRef Medline
11. Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L., Tate, J., and Punta, M. (2014) Pfam: the protein families database. *Nucleic Acids Res.* **42,** D222–D230 CrossRef Medline
12. Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3d. *Nucleic Acids Res.* **38,** W545–W549 CrossRef Medline
13. McKay, T., Hart, K., Horn, A., Kessler, H., Dodge, G., Bardhi, K., Bardhi, K., Mills, J. L., Bernstein, H. J., and Craig, P. A. (2015) Annotation of proteins of unknown function: initial enzyme results. *J. Struct. Funct. Genomics* **16,** 43–54 CrossRef Medline
14. Craig, P. (2017) BASIL: Biochemistry Authentic Scientific Inquiry Lab blog, http://basiliuse.blogspot.com
15. Craig, P. A. (2017) A survey on faculty perspectives on the transition to a biochemistry course-based undergraduate research experience laboratory. *Biochem. Mol. Biol. Educ.* **45,** 426–436 CrossRef Medline

Lessons from my undergraduate research students

Paul A. Craig

J. Biol. Chem. 2018, 293:10447-10452.
doi: 10.1074/jbc.RA118.003722 originally published online May 9, 2018